

10-23-00

A

JC943 U.S. PTO
10/20/00

NEW UNITED STATES UTILITY PATENT APPLICATION
under 37 C.F.R. 1.53(b)

Page 1

Atty. Docket No. 4910.00003

JC926 U.S. PTO
09/692884
10/20/00

CERTIFICATE OF EXPRESS MAILING UNDER 37 C.F.R. § 1.10: The undersigned hereby certifies that this United States Patent Application and all papers noted herein as being attached, are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" Service under 37 C.F.R. § 1.10 today, **October 20, 2000**, and is addressed to: Assistant Commissioner for Patents, Washington, D.C. 20231.

Express Mail Label No.: EL673059462US

Assistant Commissioner for Patents
Box Patent Applications
Washington, D.C. 20231

Enclosed herewith is a new patent application and the following papers:

First Named Inventor (or application identifier): Owens et al.

Title of Invention: Method for Establishing an MPLS Data Network Protection Pathway

■ Specification 34 pages (including specification, claims, abstract) / 16 claims (2 independent)

■ Declaration/Power of Attorney is:

■ attached in the regular manner (unexecuted).

□ NOT included, but deferred under 37 C.F.R. § 1.53(f).

■ 2 Distinct sheets of ■ Formal □ Informal Drawings

4. □ Preliminary Amendment.

5. □ Information Disclosure Statement

□ Form 1449

□ A copy of each cited prior art reference

6. □ Assignment with Cover Sheet.

7. □ Priority is hereby claimed under 35 U.S.C. § 119 based upon the following application(s):

Country	Application Number	Date of Filing (day, month, year)

8. □ Priority document(s).

NEW UNITED STATES UTILITY PATENT APPLICATION
under 37 C.F.R. 1.53(b)

Page 2

Atty. Docket No. 4910.00003

9. ☐ Statement Claiming Small Entity Status.
10. ☐ Microfiche Computer Program (Appendix).
11. ☐ Nucleotide and/or Amino Acid Sequence Submission.
☐ Computer Readable Copy.
☐ Paper Copy (identical to computer copy).
☐ Statement verifying identity of above copies.

NEW UNITED STATES UTILITY PATENT APPLICATION
under 37 C.F.R. 1.53(b)

Page 3

Atty. Docket No. 4910.00003

12. Calculation of Fees:

FEES FOR	EXCESS CLAIMS	FEE	AMOUNT DUE
Basic Filing Fee (37 C.F.R. § 1.16(a))			\$710.00
Total Claims in Excess of 20 (37 C.F.R. § 1.16(c))	0	18.00	\$0.00
Independent Claims in Excess of 3 (37 C.F.R. § 1.16(b))	0	80.00	\$0.00
Multiple Dependent Claims (37 C.F.R. § 1.16(d))	0	270.00	\$0.00
Subtotal - Filing Fee Due			\$710.00
	MULTIPLY BY		
Reduction by 50%, if Small Entity (37 C.F.R. §§ 1.9, 1.27, 1.28)	0		\$0.00
TOTAL FILING FEE DUE			\$710.00
Assignment Recordation Fee (if applicable) (37 C.F.R. § 1.21(h))	0	40.00	\$0.00
GRAND TOTAL DUE			\$710.00

13. PAYMENT is:

- ☒ included in the amount of the GRAND TOTAL by our enclosed check. A general authorization under 37 C.F.R. § 1.25(b), second sentence, is hereby given to credit or debit our Deposit Account No. 01-0850 for the instant filing and for any other fees during the pendency of this application under 37 C.F.R. §§ 1.16, 1.17 and 1.18.
- ☐ not included, but deferred under 37 C.F.R. § 1.53(f).

14. All correspondence for the attached application should be directed to:

Banner & Witcoff, Ltd.
Ten South Wacker Drive, Suite 3000
Chicago, IL 60606-7407
Telephone: (312) 715-1000
Facsimile: (312) 715-1234

15. Other: _____

Date: October 20, 2000

By: _____

Joseph P. Krause
Joseph P. Krause
Reg. No. 32,578

JPK/caj

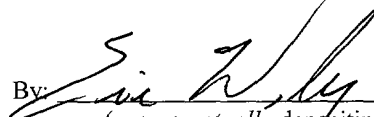
CERTIFICATE OF MAILING BY "EXPRESS MAIL"
(NEW PATENT APPLICATION)

Express Mail No. EL673059462US

Deposited October 20, 2000

JC926 U.S. PTO
09/692884
10/20/00

I hereby certify that the attached correspondence, identified below, is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" under 37 CFR § 1.10 on the date indicated above and is addressed to the Asst. Commissioner for Patents, Box Patent Application, Washington, DC 20231.

By: 
(person actually depositing)

In the Application of: Owens et al.

Serial No.: _____

Filing Date: _____

Title: Method for Establishing an MPLS Data Network Protection Pathway

- ☒ Patent Application (34 pages, 16 claims, 2 sheets formal drawings)
- ☒ Declaration and Power of Attorney (unexecuted)
- ☒ Return Postcard
- ☒ Transmittal Letter under 1.53(b) and duplicate
- ☒ Check in the amount of \$710.00

Attorney Docket No.: 4910.00003

METHOD FOR ESTABLISHING AN MPLS DATA NETWORK

PROTECTION PATHWAY

RELATED APPLICATIONS

This application claims priority under 35 U.S.C. §119(e) to provisional
5 application serial nos. 60/160,840, filed October 21, 1999; 60/161,277, filed October 25,
1999; and 60/187,798, filed March 8, 2000, the entire writing and content of which is
incorporated by reference.

FIELD OF THE INVENTION

This invention relates to data networks. In particular this invention relates to a
10 method and apparatus for automatically routing a message upstream through a multi-
protocol label switching (MPLS) network so as to trigger the re-routing of data onto an
alternate path through the MPLS network.

BACKGROUND OF THE INVENTION

Multiprotocol Label Switching (MPLS) is a new technology that combines OSI
15 layer 2 switching technologies and OSI layer 3 routing technologies. The advantages of
MPLS over other technologies include the flexible networking fabric that provides
increased performance and scalability. This includes Internet traffic engineering aspects
that include Quality of Service (QoS)/Class of Service (COS) and facilitate the use of
Virtual Private Networks (VPNs).

20 The Internet Engineering Task Force (IETF) defines MPLS as a standards-based
approach to applying label switching technology to large-scale networks. The IETF is
defining MPLS in response to numerous interrelated problems that need immediate
attention. These problems include, scaling IP networks to meet the growing demands of

Internet traffic, enabling differentiated levels of IP-based services to be provisioned, merging disparate traffic types onto a single IP network, and improving operational efficiency in a competitive environment.

The key concept in MPLS is identifying and marking IP packets with labels and forwarding them to a modified switch or router, which then uses the labels to switch the packets through the network. The labels are created and assigned to IP packets based upon the information gathered from existing IP routing protocols.

The label stack is represented as a sequence of "label stack entries". Each label stack entry is represented by 4 octets.

15

20

The label stack entries appear after the data link layer headers, but before any network layer headers. The top of the label stack appears earliest in the packet, and the bottom appears latest. The network layer packet immediately follows the label stack entry which has the S bit set.

Multi-protocol label switching (MPLS) networks are typically comprised of several packet-based switching systems interconnected by a variety of media (e.g., coaxial or fiber optic cable, unshielded twisted pair or even point-to-point microwave wireless) in a mesh-topology network similar to the public switched telephone network. In such a network, there might be several paths through the network between any two endpoints. MPLS networks carry data as packets wherein each packet includes a label on identifying a switched path through the network. The data label is appended to data packets so as to define a pathway through the network over which the data packets are to be routed.

A problem with any data network, including an MPLS network, is the amount of time required to recover from either a link failure or a switch failure. Empirical data shows that the time required to recover from a network failure can take several seconds to several minutes, an unacceptably long time. A method and apparatus by which the recovery time for a link or switch failure can be reduced to perhaps less than a few hundred milliseconds would be a significant improvement over the prior art fault recovery mechanisms used on MPLS networks to date. A method and apparatus by which a switch over from a working path to a protection path would facilitate MPLS network reliability.

SUMMARY OF THE INVENTION

In an MPLS data network comprised of various transmission media linking various types of switching systems, network fault recovery time is reduced by using a reverse-directed status message that is generated by a data switch that is down-stream from a switching system from which data is received. The reverse-directed or upstream status message is sent over a pre-determined pathway (i.e. through pre-determined

switches and/or over pre-determined data links) which originates from a destination switch or node in an MPLS network to upstream switching systems. This so-called reverse notification tree carries a message or messages that are used to indicate the functionality (or non-functionality) of the downstream switch, switches or links of the MPLS network. As long as an upstream MPLS switching system continues to receive the reverse-directed status message from a downstream switch via the reverse notification tree, the switching systems that receive such a message consider the downstream switch and pathways to be in intact. Accordingly, data packets continue to be sent downstream for subsequent routing and/or processing. If the reverse-directed status message is lost or discontinued, either because of a switch failure or a link failure, the upstream switching system considers the downstream switch or link to have failed and thereafter begins executing a procedure by which data is rerouted over an alternate data path through the network. In the preferred embodiment, the alternate data path over which downstream information is sent is a pre-established protection path and is known to a protection switch in advance, thereby minimizing data loss attributable to the time it might take to calculate a dynamic alternate protection path.

Switches in the network and their interconnections can be modeled using a directed acyclical graph by which a downstream switch knows the identity of the upstream switch to which the failure notice should be sent. In the preferred embodiment, at least one upstream switch routing the MPLS data re-directs data onto a protection path through the network between the same two endpoints by using the reverse notification tree. By way of the reverse notification tree, data loss caused by either a link or switch

failure can be minimized by the prompt rerouting of the data through an alternate or recovery data path through the network.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a simplified block diagram of an MPLS protection configuration.

5 Figure 2 depicts exemplary message flows in an MPLS network.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Figure 1 shows a simplified block diagram of a packetized-data switching network 100. Each of the squares shown in Figure 1 including boxes represented by reference numerals 102, 104, 106, 108, 110, 112, 114, 116, 118 and 120 represent one or more types of asynchronous switching systems that asynchronously receive data in e.g., packets, cells or frames from an “upstream” switch and route, direct, couple or otherwise send the data onward to another “downstream” switch logically closer to the ultimate destination for the data. By way of example, these switching systems might be internet protocol (IP) routers, asynchronous transfer mode (ATM) switches, frame relays switches or other types of packetized-data switching systems implemented to receive packetized data over a transmission line and reroute the data onto one or more output ports to which are connected transmission media coupled to other switching systems.

In Figure 1, switching system number 1, (identified by reference numeral 102) is coupled to another switching system, no. 2, (represented by reference numeral 104) and switching system no. 5 (represented by reference numeral 120) via links L_{12} and P_{25} respectively. Switching system no. 2 and no. 5 are “downstream” from no. 1; no. 1 is considered to be “upstream” from switch no. 2 and no. 5.

Similarly switching system no. 3, (represented by reference numeral 106) is coupled to switching systems 2, 4 and 9 (represented by reference numerals 104, 108 and 116 respectively) via transmission links L_{23} , L_{34} , and L_{93} respectively.

In routing data between switch no. 1 (represented by reference numeral 102) and switch no. 7 (represented by reference numeral 112) data might be routed between these two endpoints through a "*primary*" path that is comprised of links that logically or physically couple switches 2, 3, 4, 6 and 7 (identified by reference numerals 104, 106, 108, 110 and 112 respectively). The physical or logical links of the primary path between the endpoints which is 1 and 7 are represented by the vectors designated L_{12} , L_{23} , L_{34} , L_{46} and L_{67} . This path is known in the art as the working or primary path through the network. The links of the various paths shown in Figure 1 (represented by the vectors L_{12} , L_{23} , L_{34} , L_{46} and L_{67}), and therefore the paths themselves, might be constructed of direct pathways (e.g., fiber optic cable, coaxial cable, unshielded twisted pairs of copper wires, or microwave radio) between the various switches. Alternate embodiments of the paths or links between switches of the network of Figure 1 would also include using direct pathways, and intermediate switches or switch networks, (not shown in Figure 1, but still part of the path or link coupling one or more switching systems to another). By way of example and not of limitation, the data switches shown in Figure 1 might be IP switches but such IP switches could be linked together using one or more ATM switches or ATM networks.

The MPLS Protection Path

In an MPLS network, there is almost always a "*protection*" path, which is an alternate path through the network linking two endpoints. The protection path entry and

exit points are usually accessible to only *protection* switches. A protection switch is a switch that can re-route traffic onto a protection pathway. Like the other links described above, a protection pathway can be comprised of direct data paths, but also switches or switching systems, by which data can be sent through a network between two or more endpoints.

In an MPLS network, a protection path is set up using at least one protection switch element so as to be able to carry data from a source to a destination in the event the primary path or switch thereof fails for one reason or another. The operation of a protection switch is shown in Figure 1 by way of example.

In Figure 1, a working path between switch 1 and 7 exists through switches 1, 2, 3, 4, 6 and 7 and the links between the switches. A protection path for the portion of the working path that runs through switches 2, 3, 4 and 6 is the path designated by links P₂₅ and P₂₇ and which runs through switch 5, (identified by reference numeral 120). The protection path extends between endpoint switches 1 and 7 but through only switch 5 (identified by reference numeral 120). Alternate embodiments of a protection path might extend through multiple switches. In the network 100, either a link or switch loss between switch 1 and 7 can be overcome by re-routing traffic for switch 2 through switch 5 instead. Switch 5 then routes the data to switch 7. Switch 1 (identified by reference numeral 102) is considered to be a protection switch element.

Another working path between switch 8 and switch 7 of the network 100 (identified by reference numerals 114 and 112 respectively) exists through switches 9, 3, 4, 6 & 7 (identified by reference numerals 116, 106, 108 and 110 respectively) and the links between them. A protection path for data from switch 8 (reference numeral 114) to

switch 7 (reference numeral 112) through the network 100 exists via switch 10, (reference numeral 118) such that if data traffic from switch 8 (reference numeral 114) is lost somewhere between switch 9 (reference numeral 116) and switch 7 (reference numeral 112), switch 9 can re-route such data traffic to switch 10 (reference numeral 118). Switch 10 can then route such data to switch 7. Switch 9 is therefore considered to be a protection switch element.

If an alternate data path, i.e. a protection path, is pre-determined, i.e. set up or established in advance, data loss attributable to a switch or link failure can be minimized. If a protection switch is pre-programmed to re-route data upon its receipt of an appropriate command or signal, the protection switch element can almost immediately start sending data to the proper destination via the protection path.

The Liveness Message

In the event of a pathway failure, such as either a switch failure or a link failure, anywhere along a primary or working path, a protection switch element (PSL), such as switch no. 1 (identified by reference numeral 102) can re-route data traffic through the protection path so as to have the data for the endpoint switch no. 7 delivered as quickly as possible to the endpoint at switch no. 7 (identified by reference numeral 112). The ability to re-route data to a protection path is made considerably more valuable if the decision to switch over to a protection path is based upon an affirmative notice that a switch over is needed. This affirmative notice is in the form of an upstream liveness message, the loss of which indicates a pathway failure. As long as a liveness message is received at an upstream switch from a downstream switch, the upstream switch can assume that the pathway between the two switches is intact and that the downstream switch is functional.

In the event of a switch or link failure anywhere between the path endpoint switches 1 and 7, data re-routing is accomplished faster by using a reverse-directed status message that is sent backward or upstream toward the protection switch element no. 1 (reference numeral 102) by one or more of the switches 2, 3, 4, 6 or 7 (reference numerals 104, 106, 108, 110 or 112) of the primary pathway, links L_{12} , L_{23} , L_{34} , L_{46} , L_{67} . In the preferred embodiment this reverse direction data message is known as a "liveness message" the format of which is a design choice and dependent upon nature of the switches of the network 100, but the function of which is to indicate to upstream switches that the data traffic sent to the downstream switch arrived intact and on time.

The structure of a liveness message will vary depending upon whether the network switches are ATM, IP, Ethernet or other types of switches, as those skilled in the art will recognize. Unlike known fault detection methods, the liveness message is not a copy, or loop-back of the downstream data. The salient aspect of the liveness message is that it is an informational or status message, preferably sent at periodic intervals between adjacent nodes, indicating the operational condition of the switch from which it was sent. Alternate embodiments include sending a liveness message aperiodically. The fact that the liveness message is received at an upstream switch (with respect to the downstream data) is evidence that the link between the switches, over which downstream data would be sent, is intact and that the switch that generated the liveness message is at least somewhat functional.

While the preferred embodiment contemplates that the liveness message is sent upstream from a switch, directly to the switch that sent the downstream data, alternate embodiments of the invention contemplate that the liveness message could be sent

between other nodes, or from one switch to an intermediate transfer point, which for purposes of claim construction are considered to be equivalent embodiments. By way of example, with reference to Figure 1, switch no. 4 (identified by reference numeral 108) will send a liveness message, upstream to switch no. 3 (reference numeral 106) in response to data sent downstream from switch no. 3 to switch no. 4. If the liveness message from switch no. 4 is lost by or not received by switch no. 3, switch no. 3 immediately knows that either the link L_{34} between the two switches failed, or switch no. 4 failed. If switch no. 3 was carrying data from switch no. 9 (reference no. 116) and which is a protection switch element having access to a protection path, switch no. 3 would need to inhibit its liveness message to switch no. 9 or generate an error message to switch no. 9, thereby instructing switch no. 9 to re-route traffic from switch no. 3, to the protection path through switch no. 10 (reference numeral 118)

As for data routed through switch no. 3 that comes from switch no. 2 (reference numeral 104), a liveness message loss from switch no. 4 will require switch no. 3 to inhibit the liveness message to switch no. 2, or send an error message to switch no. 2. This procedure is then repeated to switch no. 1, instructing switch no. 1 to make a protection switch through switch no. 5 (reference no. 120).

Whenever the liveness message is lost, its failure is considered to be indicative of a path failure of either a link or a switch. Still other embodiments of the invention contemplate sending a downstream liveness message, sent from an upstream switch to a downstream switch thereby indicating to a downstream switch that the upstream switch and link are functional.

As set forth above, the format of a liveness message will depend upon the type of switching systems used in the network. IP switches and ATM switches will need to comply with their respective protocols. Alternative embodiments of the invention would certainly contemplate other sorts of liveness messages having different formats with the salient feature of the message being that the message indicates to an upstream switch that downstream directed data messages were received by a downstream switch intact.

In Figure 1, the links over which reverse notification status messages (i.e. the liveness messages) are sent, are designated by the reverse directed vectors, one of which is shown in Figure 1 (and identified by reference numeral R₇₆). By way of example if link L₆₇ should fail causing a data loss to the endpoint switch no. 7, the corresponding loss of the liveness message ordinarily sent from switch 7 to switch 6 would provide an indication to switch no. 6 that either the link or the switch 7 failed whereupon switch no. 6 would begin notifying the protection switch (switch no. 1) upstream from it by way of a reverse notification message R₆₄ that would be sent to switch no. 4, (represented by reference numeral 108). Similarly, switch no. 4 would thereafter return a reverse notification message R₄₃ to switch no. 3. Switch no. 3 returns another reverse notification message R₃₂ to switch 2 which then returns a reverse notification message R₂₁ to the origination node 1.

The ultimate destination of the upstream message, and in this case the reverse notification message, is a switching node (i.e. a switch or switching system) that is capable of re-routing downstream traffic, data or messages onto a different path, i.e., a protection path, usually comprised of at least a different transmission route, possibly including a different transmission media as well (coax to fiber; fiber to microwave etc.).

Whether the upstream message goes through another switch on its way to the switching node (which has the capability of re-routing data to the protection path) or is directly sent to the switching node from a downstream switch around an intermediate switch (for example, sending a liveness message directly from switch 6 to switch 1) would still
5 provide equal functionality in that the switching node will eventually receive notification that it needs to re-route traffic, data or message onto the protection path. Sending the aliveness message directly to the protection switch or routing the aliveness message via intervening switches are considered to be equivalent embodiments for purposes of claim construction.

10 Inasmuch as switch no. 1 in Figure 1 is designated as a "protection switch element" meaning that it is coupled to and capable of routing data onto a protection path P₂₅, the protection switch element 1 (identified by reference numeral 102) reroutes traffic to switch no. 7 via a protection path designated by P₂₅ and P₅₇ and that runs through switch no. 5 (identified by reference numeral 120).

15 In the preferred embodiment, the switches of the network maintain tables of network switches upon which incoming data is received and a table of network switches to which outgoing data is routed. By keeping a record of where outgoing data from a switch originates from, it is possible for a switch of the network 100 to promptly notify an upstream switch of a downstream link or switch failure.

20 In the process described above, each of the switches of the network sequentially notifies at least one switch upstream from it. Alternate (and for purposes of claim construction, equivalent) embodiments of the invention could certainly provide upstream notification messages directly from any downstream switch to every other upstream

switch in a pathway. In such an embodiment, switch no. 6 might send a reverse notification message directly to the protection switch element 1 via a direct link thereby notifying the protection switch to immediately reroute data to the protection path P₂₇ and P₅₇ via switch no. 5. Switch no. 6 might also send a reverse notification (liveness) message to the other switching systems of the network as well.

The Reverse Notification Tree

The implementation of the upstream notification message, and its conveyance upstream to a protection switch element, is enabled by using an upstream pathway denominated herein as a reverse notification tree or "RNT." The RNT is an "upstream" signal pathway that allows messages from a protection path end point to be sent "upstream" to one or more protection path (and working path) starting point switches, nodes or starting points. In the preferred embodiment, the RNT passes through the same switches and over or through the links that comprise the working path (albeit over different transmission media) and for claim construction purposes can be considered to be "coincident" with the working path. Alternate embodiments of the invention would include a reverse notification tree that runs through one or more switches or nodes that are not part of the working path, or which are only partly "coincident."

With respect to Figure 1, node 7, identified by reference numeral 112, is the RNT starting point or head end. Nodes 1 and 9, which are identified by reference numerals 102 and 116, are the end points of the RNT and to which upstream protection switch messages would be sent from any node or switch between nodes 1, 9 and 7. Intervening nodes 3, 4 and 6, identified by reference numerals 106, 108 and 110 respectively, are constituent elements or parts of the RNT.

The RNT can be established in association with the working path(s) simply by making each switching system along a working path "remember" its upstream neighbor (or the collection of upstream neighbors whose working paths converge at a network switching element and exit as one). A table or other data structure stored in memory (such as RAM, ROM, EEPROM, or disk) of the switches of the paths can be configured to

5 as RAM, ROM, EEPROM, or disk) of the switches of the paths can be configured to store data identifying switches coupled to a switching system in, or part of a working path as well as a protection path.

With respect to the network shown in Figure 1, Table 1 below shows that incoming or "Ingress" RNT messages to switch no. 3 from switch 4 are labeled "N43"

10 (not shown in Figure 1) and that these messages arrive at switch no. 3 from switch 4 at an inbound or "Ingress" interface I34 (not shown in Figure 1). Because switch no. 3 receives downstream messages from two (2) different switches, (i.e. switch 2 and switch 9) both of these two upstream switches must be sent an upstream notification therefore requiring two separate upstream messages from switch 3. Upstream RNT messages to switch 2 are

15 labeled "N32" and appear or are sent from interface I23. Upstream RNT messages to switch 9 are labeled "N93" and are sent from interface I93.

A	B	C	D	E	F
Ingress Label of RNT	Ingress Interface of RNT	Egress Label of RNT	Egress Interface of RNT	Egress Label of RNT	Egress Interface of RNT
N43	I34	N32	I23	N39	I93

20 Table 1. An inverse cross-connect reverse notification tree table for Switch 3 of Figure 1.

The reverse path (upstream) to switch 3 from switch 4 is labeled N43; the switch 3

25 interface for this data is designated I34. An upstream message received at I34 and that is

labeled N43, is sent out from switch 3, via the interfaces I23 and I93 and labeled N32 and N39 respectively.

Table 2 shows the egress and interface labels of the working or downstream path from switch 3 and the originating switches for that data.

5 The working path (downstream) path from switch 3 is to switch 4 and is labeled "L34." The switch 3 interface for this data is designated "I34." The data sent downstream from switch 3 originates from switch 2 and switch 9, which are referred to in Table 2 as "Next Hop" switches.

10 Switch no. 2 originates data to switch no. 3 and that data is received at switch no. 3 on interface "I2." Data from switch no. 9 is received at switch no. 3 at interface "I9." The RNT or upstream notification to switch no. 2 leaves switch no. 3 on its RNT interface "I23." RNT notification to switch no. 9 leaves switch no. 3 from "I93."

A	B	C	D	E	F	
Egress Label of Working Path	Egress Interface of Working Path	Next Hop IP Address of RNT	Next Hop IP Address of RNT	Egress Interface of RNT	Next Hop IP Address of RNT	Egress Interface of RNT
L34	I34	I2	I23	I9		I93

15 Table 2. An inverse cross-connect table for a hop-by-hop reverse notification tree.

20 A fault on the link between switch 3 and 4 in the downstream direction can be detected at a downstream node, switch 4 perhaps, via either a path failure (PF) or path defect (PD) condition being detected via Link Failure (LF) or Link Defect (LD) signals being propagated to an upstream switch. The downstream node will then periodically transmit fault indication signal (FIS) messages to its upstream neighbor (via the uplink R₄₃), which will propagate these further upstream (using its inverse cross-connect table) until they eventually reach the

appropriate Protection Switch Element, which will perform the protection switch. From Table 1, messages received at switch no. 3 are labeled "N43." Therefore, in Fig. 1, if link L34 has a fault, switch 3 will detect the fault via the lost liveness message from switch no. 4 and start transmitting an FIS packet back to switch 2 link L₂₃ as represented by the message R₃₂. From Tables 1 and 2, there are two egress messages and interfaces from switch no. 3, which identify the upstream switches that are to be "notified" of a failure downstream from switch no. 3. (The traffic in the queues of switch 3 will continue to be serviced.) By using similar tables, switch 2 in turn will propagate the FIS over the RNT back to switch 1. The actual protection switch will be performed by switch 1, after the receipt of the first FIS. Switch 3 will stop transmitting FIS messages "t" time units after the transmission of the first FIS message.

In the preferred embodiment, only one RNT is required for all the working paths that merge (either physically or virtually) to form the multipoint-to-point "forward" or "downstream" path. Figure 1 shows that at least two (2) working paths (one path of which is comprised of switch elements 1, 2 and 3 that are identified by reference numerals 102, 104 and 106; a second path of which is comprised of switch elements 8, 9 and 3 that are identified by reference numerals 114, 116 and 106) converge at switch element 3 (identified by reference numeral 106). Alternate (and for purposes of claim construction, equivalent) embodiments would include using multiple RNTs for a single working path that has multiple paths that converge at a single node (switches of each path that converges might form different RNTs) as well as using multiple RNTs for a single working path.

The RNT is rooted at an appropriately chosen label switched router ("LSR"), (which hereafter is referred to as an MPLS network switch element) along the common segment of the merged working paths and is terminated at the protection switch elements (PSLs). Intermediate network switching elements on the converged working paths typically share the same RNT reducing signaling overhead associated with recovery.

Unlike schemes that treat each network switch element independently, and require signaling between a protection switch element and a destination switch individually for each network switch element, the RNT allows for only one (or a small number of) signaling messages on the shared segments of the label switch paths (LSPs).

The RNT can be implemented either at Layer 3 or at Layer 2 of the OSI, 7-layer protocol stack. In either case, delay along the RNT needs to be carefully controlled. This may be accomplished by giving the highest priority to the fault and repair notification packets, which travel along the RNT. We can therefore have a situation where different protection domains share a common RNT.

A protection "domain" is considered to be the switches and links of both a working path and protection path. For example, in Fig. 1, the protection domain bounded by network switch element 1 and network switch element 7, is denoted by {1-2-3-4-6-7, 1-5-7}.

When different protection domains have different RNTs, two cases may arise, depending on whether or not any portions of the two domains overlap, that is, have nodes or links in common. If the protection domains do not overlap, the protection domains are considered to be independent. By virtue of the RNTs in the two domains being different, neither of the working paths nor the RNTs of the two domains can overlap. In other

words, failures in one domain do not interact with failures in the other domain. For example, the protection domain defined by {9-3-4- 6-7, 9-10-7} is completely independent of the domain defined by {11-13-5-15, 11-13-14-15}. As a result, as long as faults occur in independent domains, the network shown in Fig. 1 can tolerate multiple faults (for example, simultaneous failures on the working path in each domain). If protection domains with disjoint RNTs overlap, it implies that the protection path of one intersects the working path of the other. Therefore, although failures on the working paths of the two domains do not affect one another, failures on the protection path of one may affect the working path of the other and visa versa. For example, the protection domain defined by {1-2-3-4-6-7, 1-5-7} is not independent of the domain defined by {11-13-5-15, 11-13-14-15} since LSR 5 lies on the protection path in the former domain and on the working path in the latter domain. When protection domains have the same RNT, different failures along the working paths may affect both paths differently. As shown in Fig. 1, for example, working paths 1-2-3-4-5-7 and 9-3-4-6-7 share the same RNT. As a result, for a failure on some segments of the working path, both domains will be affected, resulting in a protection switch in both (for example, the segment 3-4-6-7 in Fig. 1).

Likewise, for failures on other segments of the working path, only one domain may be affected (for example, failure on segment 2-3 affects only the first working path 1-2-3-4-6-7, where as failure on the segment 9-3 affects only the second working path 9-3-4-6-7).

There are a number of ways to establish a protection domain, i.e., a working path and a protection path through an MPLS network. Establishing a protection path first requires the identification of the working path (embodied as some series of switches and path links through the MPLS network from a sending node to a destination node). In most

cases, the working path and its corresponding recovery path are specified during a network switch path or connection setup procedure, either via a path selection algorithm (running at a centralized location or at an ingress network switch element) or via an administrative configuration (e.g. a manual specification of switches that comprise the protection path).

The specification of either a protection or working path, does not, strictly speaking, require the entire path to be explicitly specified. Rather, it requires only that the head end node or switching node and end or destination switch or node (of the respective paths) be specified. In the absence of a destination switch/node specification, the path egress points out of the MPLS network or domain need to be specified, with the segments between them being – “loosely” determined or routed. In other words, a working path would be established between the two nodes at the boundaries of a protection domain via (possibly loose) explicit (or source) routing using LDP/RSVP [label distribution protocol/reservation protocol] signaling (alternatively, via constraint-based routing, or using manual configuration), as set forth more fully below.

Figure 2 depicts message flows between four (4) different switches of an MPLS network that employs the path protection techniques disclosed herein. Vertical axes of Figure 2, identified by reference numerals 202, 204, 206 and 208, represent switching elements (shown in Figure 1) of an MPLS network from which and to which various types of messages are received and sent respectively. Switch 202 is upstream from switches 204, 206 and 208. Switch 204 is upstream from switch 206 as switch 206 is upstream from switch 208.

Protection Path Establishment

A Protection Domain Path is established by the identification of a protection switch or node and an end point switch or node in the MPLS network. The protection switch element ("PSL") initiates the working network switch elements and the recovery network switch element. It is also responsible for storing information about which network switch elements or portions thereof have protection enabled, and for maintaining a binding between outgoing labels specifying the working path and the protection/recovery path. The latter enables the switchover to the recovery path upon the receipt of a protection switch trigger.

A "label distribution protocol" is a set of procedures by which one LSR (i.e., a network switch element) informs another of the label bindings it has made. "Label binding" is a process by which a message to be sent from a source to a destination is associated with various labels between the nodes that lie along the way, between the source and destination. By way of example, in Figure 1, a message to be sent from switch 1 to switch 7 is associated or bound to travel to switch 7 through switch 2 by, or using, the label L_{12} that is first associated with the message at, or by, switch 1. Switch 2 in turn associates messages labeled L_{12} as bound for switch 3 and re-labels them as L_{23} . Re-labeling messages (e.g. re-labeling a message received at switch 2 on L_{12} , as the same message that is output from switch 2 but on L_{23} and which is received at switch 3, to be re-labeled by switch 3 and output again as L_{34}) is known as "label binding." Two or more LSRs, (network switch elements) which use a label distribution protocol to exchange label binding information are known as "label distribution peers" with respect to the binding information they exchange.

The label distribution protocol also encompasses any negotiations in which two, label distribution peers, need to engage in order to learn of each other's MPLS capabilities. This label distribution protocol is referred to as path establishment signaling. MPLS defines two methods for label distribution. These two methods are: Label

5 Distribution Protocol (LDP/CR-LDP) and ReSerVation Protocol (RSVP).

Both LDP/CR-LDP and RSVP allow a path to be setup loosely (wherein each node determines it's next hop) or explicitly (wherein each node has been given it's next hop). These two protocols can be extended, as disclosed herein and by equivalents thereof, to provide a novel mechanism by which protection path establishment can be signaled and created. Accordingly, a "Protection" field can be defined, and added as an extension to the existing label request messages in LDP/CR-LDP, and path message in RSVP protocols. The destination or end point node in the MPLS network participates in setting up a recovery path as a merging network switch element. The destination or end point node learns, during a signaling or working/protection path configuration process, which working and protection paths are merged to the same outgoing network switch element.

Hosts and routers that support both RSVP and Multi-Protocol Label Switching can associate labels with RSVP flows. When MPLS and RSVP are combined, the definition of a flow can be made more flexible. Once a label switched path (LSP) is established, the traffic through the path is defined by the label applied at the ingress node of the LSP (label switched path). The mapping of a label to traffic can be accomplished using a number of different criteria. The set of packets that are assigned the same label value by a specific node are said to belong to the same forwarding equivalence class

(FEC) and effectively define the "RSVP flow." When traffic is mapped onto a label-switched path in this way, we call the LSP an "LSP Tunnel". When labels are associated with traffic flows, it becomes possible for a router to identify the appropriate reservation state for a packet based on the packet's label value.

5 A Path message travels from a sender to receiver(s) along the same path(s) used by the data packets. The IP source address of a Path message must be an address of the sender it describes, while the destination address must be the DestAddress for the session. These addresses assure that the message will be correctly routed through a non-RSVP cloud.

10 The format of an exemplary RSVP message with the Protection Object extension is:

```

15           <Path Message> ::= <Common Header> [ <INTEGRITY> ]
                  <SESSION> <RSVP_HOP>
                  [ <TIME_VALUES> ]
                  [ <EXPLICIT_ROUTE> ]
                  [ <PROTECTION> ]           /* The new message field. */
                  <LABEL_REQUEST>
20           [ <SESSION_ATTRIBUTE> ]
                  [ <POLICY_DATA> ... ]
                  <sender descriptor>

```

25 Label Distribution Protocol (LDP) is defined for distribution of labels inside one MPLS domain. One of the most important services that may be offered using MPLS in general, and LDP in particular, is support for constraint-based routing of traffic across the routed network. Constraint-based routing offers the opportunity to extend the information used to setup paths beyond what is available for the routing protocol. For instance, an

[illegible]

These requirements may be met by extending LDP for support of constraint-based routed

The Path Vector TLV is used with the Hop Count TLV in Label Request and Label Mapping messages to implement the optional LDP loop detection mechanism. Its use in the Label Request message records the path of LSRs the request has traversed. Its use in the Label Mapping message records the path of LSRs a label advertisement has traversed to setup an LSP.

15

20

25

30

35

23

Wherein the “Protection TLV” message field is new.

The Protection Object (RSVP)/Protection Type Length Value (TLV) (LDP/CR-

5 LDP) establishes the working and a corresponding protection path utilizing the
Reservation Protocol (RSVP) path message or the Constraint-Based Routing Label
Distribution Protocol (CR-LDP) Label Request message. The attributes required to
establish the Protection Domain are:

10

1 Priority: Specifies whether this protection group is a high or low switching priority.

2 Protection ID: Specifies whether protection is supported.

3 Protection Type: Specifies whether this establishment is for the Protection, or
Working Path.

15

4 Protection Identity: Specifies a unique identifier for the protection traffic.

5 Node Identity: Specifies whether the node is a switching, merging, or RNT root node.

6 RNT Type: Specifies whether the RNT is created using Hop-by-hop, MPLS LSP, or
SONET K1/K2.

7 Timer Options: Specifies the hold off and notification time requirements.

20

8 Recovery Option: Specifies whether the recovery is revertive and if the action is Wait,
Switch Back, or Switchover.

9 Protection Bandwidth: Specifies whether the bandwidth of the protection path is
available to carry excess (preemptable) traffic.

The following table illustrates the structure of an exemplary Protection

Object/Protection TLV Structure.

5	0	1	2	3																				
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1		
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	
	P	D	T	PGID		NID	RNTT	TO	RO	B	RESVD													
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	

10

P= Priority

D= Protection ID

T= Protection Type

PGID= Protection path Identity

15

NID=Node Identity

RNTT= RNT Type

TO=Timer Options

RO=Recovery Option

B= Protection Bandwidth

20

RESVD=Reserved for Future Use

Since the switching systems used in the network 100 are unidirectional, and pathway fault recovery requires the notification of faults to a protection switch, such as switch no. 1 or switch no. 9, responsible for a switchover to a recovery path, a mechanism is provided for the fault indication and the fault recovery notification to travel from a point of occurrence of the fault back to the protection switch. The ability to propagate a fault notice upstream however is complicated when two or more data streams merge in a single switch such as the streams from switches 9 and 2 merging at switch 3. When two or more data streams merge at a switch, e.g. switch 9, a fault anywhere downstream from switch 9 will require that a fault notice be sent to multiple source switches, i.e. switches 9 and 2. The fault indication and recovery notification should be able to travel along a reverse path of the working paths to all

the protection switch elements that might be affected by the fault. The path is provided by the reverse notification tree.

The MPLS protection switch message sequence begins with the establishment of the particular working paths and protection paths through the network.

- 5 The establishment of the working path and protection path is accomplished by the transmission of a Protection Switch Domain (PSD) initialization message 210 from a switch 202 to switches 204, 206 and 208. A PSD confirmation message 212 is propagated from the downstream switch 208 upstream to switch 202.

- 10 The reverse notification path, also referred to above as the Reverse Notification Tree or RNT, is established by the downstream switch, 208, sending an RNT initialization message 214, upstream to switches 206, 204 and 202. Confirmation of the RNT setup is accomplished by the RNT Confirmation message 216 that originates from switch 202.

- 15 Upon the establishment of the working and protection paths, and the reverse notification tree, data 218 can be sent through the network.

- 20 Two “aliveness” messages 220 and 222, which provide notification of the working path status, are shown in Figure 2 to depict the fact that the aliveness message described above can be sent periodically, regardless of whether downstream data 218 was sent. As shown further, downstream data transmissions, such as transmissions 224, 226 and 228 are not conditioned upon receipt of an aliveness message in any fixed way. An aliveness message 230 sent upstream is then followed by yet another data transmission 232.

Figure 2 shows that the sequence of aliveness messages and data transmissions do not need to follow any sort of predetermined or fixed order. For network reliability purposes, the aliveness messages are preferably sent periodically, so that their absence can be detected if they do not arrive on time. Alternate embodiments include sending

5 liveness messages aperiodically.

Those skilled in the art will recognize that re-routing data on a either the failure of a link or a switch in a network such as that depicted in Figure 1 need not be performed by a protection switch. In the event that switch 4 fails for example switch no. 3 might reroute data from switch 2 that is destined for switch 7, through another protection switch

10 element 9, identified by reference numeral 116. Switch 9 might then reroute data from switch 2 that is addressed to switch 7 over a protection path designed as $P_{9,10}$ and $P_{10,7}$ through switch 10, identified as reference numeral 118.

In the preferred embodiment, the media over which data message are carried might be twisted copper wires, coax cable, fiber optic cable or even a radio frequency

15 data link. As set forth above, each of the switching systems might accommodate a variety of packetized data messages including but not limited to Ethernet, internet protocol, ATM, frame relay or other types of transmission switching systems.

By continuously sending an upstream message indicating that downstream traffic arrives at its destination, recovery time required to recover from the fault of a media link

20 or a switching system can be minimized. If the switch status message used to indicate a functionality of a switch or a link is sent promptly enough, and to the appropriate node in a mesh network such as that shown in Figure 1, the time required to reroute data messages between first and second endpoint switches over an alternate data path can be

minimized. In the preferred embodiment, the alternate or so called protection path is preferably set up in advance and maintained in a stand by mode such that it is immediately available when required by the protection switch that will reroute data over the protection path.

11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 26

We claim:

- 1 1. In a multi-protocol label switching system (MPLS) data network comprised of
2 a plurality of data switches that are interconnected to form a plurality of data
3 paths from a source node to a destination node through a first set of data
4 switches, a method of establishing a protection path from said source switch to
5 said destination switch through a second set of switches, said method
6 comprised of the steps of:
7 a. sending a first predetermined control message, over a first data path from
8 said first switch to a second switch, said first predetermined message
9 establishing at least a working path through said network between said
10 first and second switches.
11 b. sending a second predetermined control message, over a second data path
12 from said first switch to a second switch, said second predetermined
13 message establishing at least a protection path through said network
14 between said first and second switches.
15 c. Associating said first working path to said protection path to enable
16 protection switching.
- 1 2. The method of claim 1 wherein said step of sending at least a first
2 predetermined message is comprised of the step of adding a protection
3 messaging field to a label distribution protocol (LDP) message, said protection
4 messaging field carrying protection pathway information between MPLS
5 network switch elements.

1 3. The method of claim 1 wherein said step of sending at least a first
2 predetermined message is comprised of the step of adding a protection
3 messaging field in an MPLS reservation protocol message (RSVP), said
4 protection field carrying protection pathway information between MPLS
5 network switch elements.

1 4. The method of claim 1 wherein said step of sending at least a first
2 predetermined message, over a first data path from said first switch to a
3 second switch, said first predetermined message establishing at least a
4 working path through said network between said first and second switches
5 includes the step of:
6 a. identifying at least one data switch of said MPLS network as a switch
7 element by the contents of at least one control field in a message field of
8 an MPLS message;
9 b. sending said at least one control field to at least one data switch of said
10 MPLS network.

1 5. The method of claim 1 wherein said step of sending at least a first
2 predetermined message, over a first data path from said first switch to a
3 second switch, said first predetermined message establishing at least a
4 protection path through said network between said first and second switches
5 includes the step of:

- 6 a. identifying at least one data switch of said MPLS network as a protection
7 switch element by the contents of at least one control field in a message
8 field of an MPLS message;
9 b. sending said at least one control field to at least one data switch of said
10 MPLS network.

1 6. The method of claim 1 wherein said step of label binding said first
2 predetermined message from said second switch to a third switch.

1 7. The method of claim 1 wherein said first data path is set up loosely.

1 8. The method of claim 1 wherein said first data path is set up explicitly.

1 9. The method of claim 1 further including the step of mapping labels to data
2 routed along said first data path according to predetermined criteria that
3 includes the quality of service to be granted said data.

1 10. In a multi-protocol label switching system (MPLS) data network comprised of
2 a plurality of data switches that are interconnected to form a plurality of data
3 paths from a source node to a destination node through a first set of data
4 switches, a method of establishing a working path from said source switch to
5 said destination switch through said first set of switches, said method
6 comprised of the steps of:

7 a. sending at least a first predetermined control message, over a first control
8 path from said first switch to a second switch, said first predetermined
9 control message establishing at least a working path through said network
10 between said first and second switches over which data is to be sent from
11 said source switch to said destination switch.

1 11. The method of claim 10 wherein said step of sending at least a first
2 predetermined control message is comprised of the step of adding a protection
3 messaging field to a label distribution protocol (LDP) message, said protection
4 messaging field carrying protection pathway information between MPLS
5 network switch elements.

1 12. The method of claim 10 wherein said step of sending at least a first
2 predetermined control message is comprised of the step of adding a protection
3 messaging field in an MPLS reservation protocol message (RSVP), said
4 protection field carrying protection pathway information between MPLS
5 network switch elements.

1 13. The method of claim 10 wherein said step of sending at least a first
2 predetermined control message, over a first data path from said first switch to
3 a second switch, said first predetermined control message establishing at least
4 a protection path through said network between said first and second switches
5 includes the step of:

- 6 a. identifying at least one data switch of said MPLS network as a protection
7 switch element by the contents of at least one data field in a message field
8 of an MPLS message;
9 b. sending said at least one data field to at least one data switch of said
10 MPLS network.

1 14. The method of claim 10 wherein said first data path is set up loosely.

1 15. The method of claim 10 wherein said first data path is set up explicitly.

1 16. The method of claim 10 further including the step of mapping labels to data
2 routed along said first control path according to predetermined criteria that
3 includes the quality of service to be granted said data.

ABSTRACT OF THE DISCLOSURE

Recovery time upon the failure of a link or switching system in an asynchronous data network can be minimized if downstream data switches provide upstream messages indicating to upstream switching system that the downstream traffic arrived in tact and was properly handled. Upon this loss or failure of the upstream status message to an upstream switching system, an upstream switching system can reroute data traffic around a failed link or failed switch with a minimal amount of lost data. The upstream status message is conveyed from a downstream switching system to an upstream switching system via a reverse notification tree data pathway. The reverse notification pathway is established by using a messaging protocol between network switches by which the working and protection paths, among other things, are established.

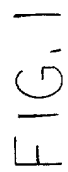
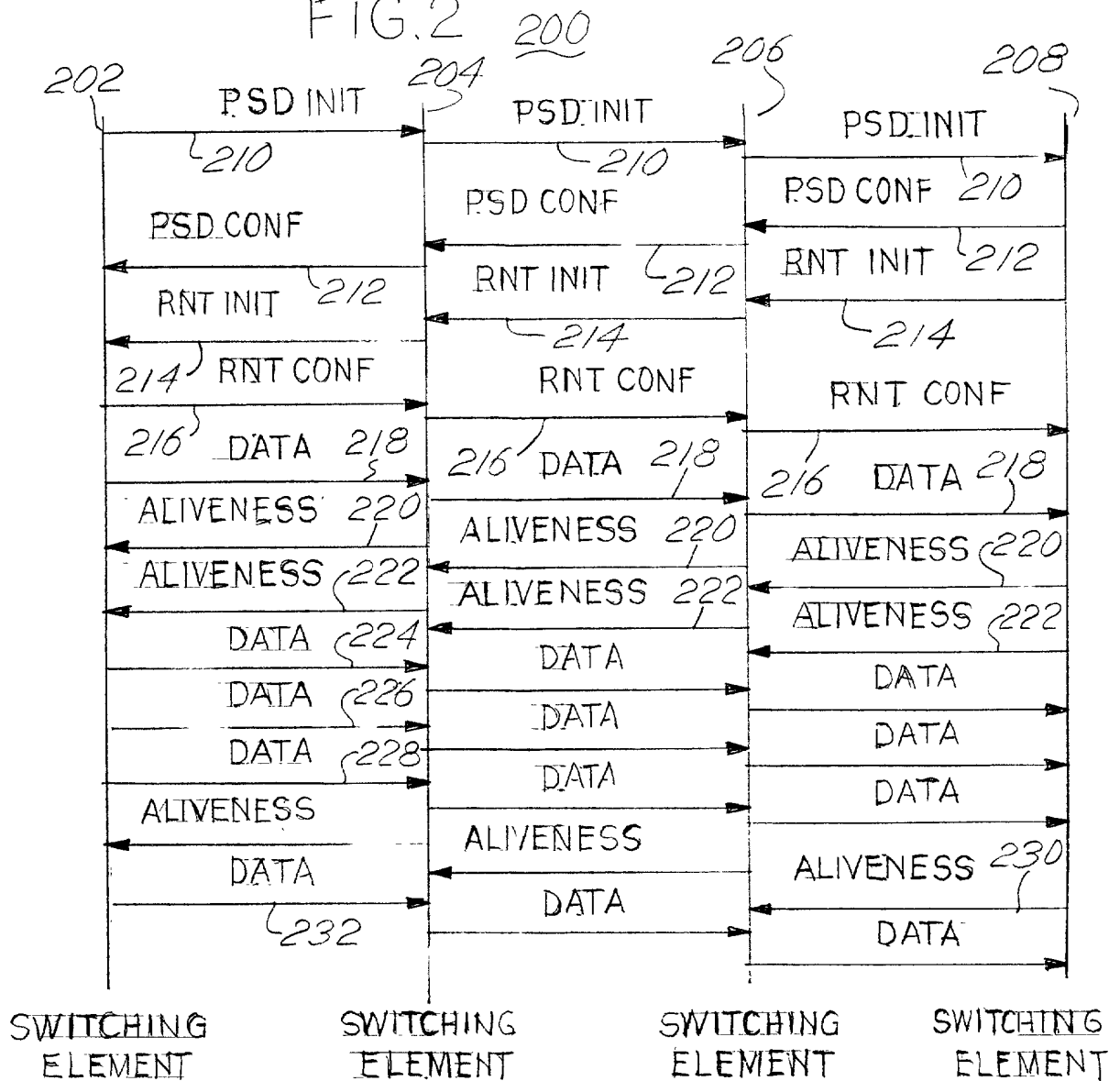


FIG. 2



PSD=PROTECTION SWITCH DOMAIN(ESTABLISHMENT OF THE WORKING AND PROTECTION PATH: IDENTIFICATION OF SWITCHING ELEMENTS)
 RNT=REVERSE NOTIFICATION TREE (ESTABLISHMENT OF REVERSE PATH FOR NOTIFICATION OF ALIVENESS)
 INIT=INITIALIZE
 CONF=CONFIRM
 DATA=DATA FLOW
 ALIVENESS=NOTIFICATION OF WORKING PATH STATUS